Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search?

Abstract

Systematic reviews are comprehensive literature reviews for a highly focused research question. These reviews are considered the highest form of evidence in medicine. Complex Boolean queries are developed as part of the systematic review creation process to retrieve literature, as they permit reproducibility and understandability. However, it is difficult and time-consuming to develop high-quality Boolean queries, often requiring the expertise of expert searchers like librarians. Recent advances in transformer-based generative models have shown their ability to effectively follow user instructions and generate answers based on these instructions. In this paper, we investigate ChatGPT as a means for automatically formulating and refining complex Boolean queries for systematic review literature search. Overall, our research finds that ChatGPT has the potential to generate effective Boolean queries. The ability of ChatGPT to follow complex instructions and generate highly precise queries makes it a tool of potential value for researchers conducting systematic reviews, particularly for rapid reviews where time is a constraint and where one can trade off higher precision for lower recall. We also identify several caveats in using ChatGPT for this task, highlighting that this technology needs further validation before it is suitable for widespread uptake.